

# Covariant Lyapunov vectors

**Francesco Ginelli**

SUPA, Institute for Complex Systems and Mathematical Biology, King's College, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom

**Hugues Chaté**

Service de Physique de l'État Condensé, CEA-Saclay, 91191 Gif-sur-Yvette, France

**Roberto Livi**

Dipartimento di Fisica, INFN and CSDC, Università di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Italy and ISC-CNR, via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy

**Antonio Politi**

SUPA, Institute for Complex Systems and Mathematical Biology, King's College, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom

**Abstract.** The recent years have witnessed a growing interest for covariant Lyapunov vectors (CLVs) which span local intrinsic directions in the phase space of chaotic systems. Here we review the basic results of ergodic theory, with a specific reference to the implications of Oseledets' theorem for the properties of the CLVs. We then present a detailed description of a “dynamical” algorithm to compute the CLVs and show that it generically converges exponentially in time. We also discuss its numerical performance and compare it with other algorithms presented in literature. We finally illustrate how CLVs can be used to quantify deviations from hyperbolicity with reference to a dissipative system (a chain of Hénon maps) and a Hamiltonian model (a Fermi-Pasta-Ulam chain).

PACS numbers: 05.70.Ln, 05.90.+m, 45.70.-n, 87.18.Ed

## 1. Introduction

Lyapunov exponents (LEs) have been recognized as a central tool for the characterization of chaotic dynamics since the late '70s when effective algorithms have been independently proposed by Shimada and Nagashima [1] and Benettin et al. [2]. In fact, although it had been theoretically clarified that LEs provide the right framework for an appropriate (coordinate-independent) characterization of dynamical regimes (see the pioneering work by Oseledets [3]), it was necessary to wait until an effective algorithm was made available to a wide community of scientists, to appreciate the usefulness of such a tool in physically relevant models, where it is not obvious whether theoretical results apply. It is nowadays widely recognized that LEs help to quantify a number of interesting physical properties such as dynamical entropies and fractal dimensions [4].

Since the LEs can be viewed as the eigenvalues of suitable matrices, one might have expected a comparable interest to be devoted to the corresponding Lyapunov vectors. Unfortunately, for a long time, the only numerically accessible vectors were those ones obtained as a by-product of the Gram-Schmidt orthogonalization procedure (an intrinsic step of the algorithms defined in [1, 2]), which are coordinate-dependent. As a result, Lyapunov vectors did not attract much interest within the nonlinear-dynamics community, even though an objective definition of intrinsic, covariant Lyapunov vectors<sup>†</sup> (CLVs) has been given three decades ago [5]. In fact, until recently, as a result of the lack of effective algorithms, only a few papers can be found in the literature where the proper CLVs have been determined and used to characterize chaotic states [6, 7, 8].

The situation has drastically changed a few years ago, when efficient algorithms have been developed [9, 10, 11], so that many scientists are now aware that CLVs can offer information on the local geometric structure of chaotic attractors, as opposed to LEs, which are powerful but global quantities. Several papers appeared, where CLVs have been successfully employed to better understand many aspects of chaotic dynamics [12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Some of the relevant questions are touched in this Special Issue.

The goal of this paper is to provide a fairly general introduction to the topic, including a description of the various algorithms that have been proposed, but mostly focusing on the “dynamical” approach that we believe to be the most effective one [9]. In particular, we discuss the convergence properties of the CLVs by making use of a suitable large-deviation function. An additional point that we explore is the minimal angle between the unstable and stable manifold, as it provides direct information on the hyperbolicity of the underlying dynamics. In particular, we find that angle distribution satisfies a natural scaling law, both in a dissipative and Hamiltonian context.

More precisely, we start with a discussion of the theoretical context where CLVs can be properly defined. This is done in Section 2, which contains a short review of the basic results of ergodic theory in the generic context of smooth Riemannian manifolds. In Section 3, we then review some implications of Oseledets’ theorem in the more practical context of  $\mathbb{R}^N$  and discrete-time dynamics (to which every continuous-time dynamics has to be reduced in computer simulations) for time-reversible systems. In Section 4, we introduce covariant Lyapunov vectors, describe in

<sup>†</sup>Sometimes they are also referred as *characteristic* Lyapunov vectors.

detail the algorithm introduced in [9], and discuss its application to non-invertible dynamics as well. The convergence properties of the algorithm are discussed in Section 5, with reference to coupled Hénon maps and Fermi-Pasta-Ulam chains. In Section 6 we briefly review other algorithms that have been proposed in the literature and mutually compare them. Section 7 is devoted to a discussion of the angles between CLVs, essentially focussing on the angle between the stable and unstable manifold in a dissipative (coupled Hénon maps) and Hamiltonian (Fermi-Pasta-Ulam chain) model. We conclude by briefly mentioning other relevant applications and pending questions.

## 2. Basic tools from ergodic theory

The mathematical foundations of ergodic theory provide a solid basis for the study of dynamical systems and, in particular, of their stability properties. For what concerns the application to various fields like statistical and celestial mechanics, plasma and accelerator physics, fluid dynamics, meteorology etc., the main contribution of ergodic theory stems from the possibility of translating rigorous mathematical results into effective algorithms, that allow for a quantitative analysis. It is worth to point out that the tools of ergodic theory apply to both continuous (in time) as well as to discrete (in time) dynamical systems, i.e. flows and maps, respectively. The study of the Lyapunov characteristic exponents, or simply Lyapunov exponents, is one of the main chapters of ergodic theory and goes back to the beginning of the XX-th century, when the Russian mathematician Aleksandr M. Lyapunov introduced them as indicators of the stability of a singular point or of a periodic orbit. As mentioned in the introduction, here we aim at providing an overview about covariant Lyapunov vectors and their practical use in applications. Accordingly, in this section we summarize the main results of ergodic theory that are strictly related to the concept of CLV and to the possibility of explicitly computing them ‡.

### 2.1. Dynamical systems, stability and measures

Oseledets' theorem is a major achievement in the study of the stability of dynamical systems. It proves the existence of LE for generic orbits under quite general conditions [3]. In practice, this theorem offers the possibility of extending Lyapunov stability analysis from fixed points and periodic orbits, to any trajectory of a dynamical system  $(\mathcal{M}, L)$ , either continuous or discrete in time, defined on a Riemannian manifold  $\mathcal{M}$  of dimension  $N$  and equipped with a suitable metric. The application  $L$  is the evolution operator of the dynamical system, i.e.  $L^t(x) \equiv x_t$ , where  $x_t$  is the image at time  $t$  of the initial condition  $x$  at time  $t = 0$ . We assume that at each point  $x \in \mathcal{M}$  we can identify a tangent space  $\mathcal{T}_x\mathcal{M}$ . With reference to a vector  $\xi \in \mathcal{T}_x\mathcal{M}$  we can define the expansion rate

$$\gamma(\xi, x, t) = \frac{\|DL^t(x)\xi\|}{\|\xi\|} \quad (1)$$

where the symbol  $\|\bullet\|$  is the norm and  $DL^t(x) : \mathcal{T}_x\mathcal{M} \rightarrow \mathcal{T}_{L^t(x)}\mathcal{M}$  indicates the evolution in tangent space. The Lyapunov characteristic exponent associated to the

‡For a more general survey about ergodic theory we suggest the contribution by Lai-Sang Young in this special issue

vector  $\xi \in \mathcal{T}_x\mathcal{M}$  is defined as

$$\lambda(\xi, x) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\|DL^t(x)\xi\|}{\|\xi\|} \quad (2)$$

provided this limit exists. This quantity essentially tells us how an infinitesimal perturbation  $\delta x$  of the initial condition  $x \in \mathcal{M}$  along the vector  $\xi \in \mathcal{T}_x\mathcal{M}$  is exponentially expanded (or contracted) at large times

$$\|\delta x_t\| \sim \|\delta x\| e^{\lambda(\xi, x)t}. \quad (3)$$

The crucial point here is that for a generic (non periodic) trajectory in  $\mathcal{M}$  at any  $t$  the tangent application  $DL^t(x)$  applies to different tangent spaces at different times, and the very notion of eigenvalues and eigenvectors of this application is undefined, even if the limit (2) exists. This notwithstanding, a filtration, i.e. a local decomposition into suitable subspaces of  $\mathcal{T}_x\mathcal{M}$  at any  $x \in \mathcal{M}$ , still exists as in the case of fixed points and periodic orbits. In fact, when the vector  $\xi$  is varied in  $\mathcal{T}_x\mathcal{M}$  the quantity  $\lambda(\xi, x)$  takes a finite number  $m \leq N$  of distinct values  $\lambda_1^*(x) > \lambda_2^*(x) > \dots > \lambda_m^*(x)$ . The filtration of the tangent space into subspaces  $S_i$ ,  $\mathcal{T}_x\mathcal{M} \equiv S_1 \supset S_2 \supset \dots \supset S_m$ , is such that by choosing  $\xi \in S_i \setminus S_{i+1}$ , one has  $\lambda(\xi, x) = \lambda_i^*(x)$ . It is important to point out that in general one can associate to each  $\lambda_i^*$  a multiplicity (or degeneracy)  $g_i = \dim S_i - \dim S_{i+1}$ . A basis of  $\mathcal{T}_x\mathcal{M}$ , obtained by taking in each subspace  $S_i \setminus S_{i+1}$  a number  $g_i$  of independent vectors is called a *normal basis*. It can be easily argued that if  $(f_1, f_2, \dots, f_N)$  is a generic basis of the tangent space  $\mathcal{T}_x\mathcal{M}$  and  $(e_1, e_2, \dots, e_N)$  is a normal basis one has

$$\sum_{i=1}^N \lambda(e_i, x) \leq \sum_{i=1}^N \lambda(f_i, x) \quad (4)$$

where the equality holds only if  $(f_1, f_2, \dots, f_N)$  is a normal basis. The ordered sequence of the  $m$  characteristic exponents  $\lambda_i^*(x)$ , each one repeated with its multiplicity  $g_i$  is called the spectrum,  $Sp(x)$ , of characteristic Lyapunov exponents  $\lambda_1(x) \geq \lambda_2(x) \geq \dots \geq \lambda_N(x)$ . In practice, the existence of a filtration amounts to establishing the existence of linear subspaces  $E_i = S_i \setminus S_{i+1}$  of dimension  $g_i$  that identify the characteristic Lyapunov exponent  $\lambda_i(x)$  together with its multiplicity. One can easily realize that the extension of such a concept to linear subspaces  $E \subset \mathcal{T}_x\mathcal{M}$  with  $\dim E = p \leq N$  is straightforward. In fact one can define the Lyapunov characteristic exponent of order  $p$

$$\lambda(E, x) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\text{Vol}^p(DL^t(x)\xi_1, \dots, DL^t(x)\xi_p)}{\text{Vol}^p(\xi_1, \dots, \xi_p)} \quad (5)$$

provided this limit exists, where  $\text{Vol}^p(\dots)$  is the volume of the  $p$ -dimensional parallelepiped generated by the tangent vectors in its argument. Let us stress that the LE defined in equations (2) and (5) are independent of the chosen metric.

After all of these preliminary considerations one can summarize the very content of Oseledets' theorem that applies to a dynamical system  $(\mathcal{M}, \mu, L)$  that preserves the measure  $\mu$  and can be either continuous or discrete in time, invertible or non-invertible. The theorem states that [3]

*For almost all  $x \in \mathcal{M}$  and for any subspace  $E \subset \mathcal{T}_x\mathcal{M}$ , such that  $\dim E = p \leq N$  the limit in Eq. (5) exists and is finite and, in particular, the limit in Eq. (2) exists and is finite for any tangent vector  $\xi \in \mathcal{T}_x\mathcal{M}$ . Moreover the*

spectrum is a measurable function of  $x$  and for any subspace  $E \subset \mathcal{T}_x\mathcal{M}$  it exists a normal basis  $(e_1, e_2, \dots, e_N)$  such that

$$\lambda^{(p)}(E, x) = \sum_{i=1}^p \lambda(e_i, x). \quad (6)$$

Some interesting consequences of this theorem are worth to be listed. First of all, one can easily realize that LE are constant of the motion generated by the evolution operator  $L$ , thus implying that for an ergodic system they are almost-everywhere constant,  $\lambda_i(x) = \lambda_i$ . If the dynamical system is invertible and the measure  $\mu$  is conserved, the sum of the Lyapunov characteristics exponents (or the Lyapunov exponent of order  $N$ ) is zero,  $\lambda^{(N)}(\mathcal{T}_x\mathcal{M}, x) = \sum_{i=1}^N \lambda(e_i, x) = 0$ . In particular, for a Hamiltonian system and, more generally, for any symplectic diffeomorphism (i.e., canonical transformation of the manifold  $\mathcal{M}$  onto itself) the spectrum of LE is symmetric, i.e.  $\lambda_i(x) = -\lambda_{N-i+1}(x)$  for  $i = 1, \dots, N/2$ , where each individual exponent  $\lambda_{N/2-j}(x) = 0$  for  $j = 0, \dots, n_c - 1$  if there are  $n_c$  constants of the motion (one of which is necessarily the energy of the system). Moreover, for any continuous time dynamical system whose support does not reduce to a fixed point, at least one of the Lyapunov exponents should vanish.

Finally, it is worth pointing out one further major consequence of Oseledets' theorem that is of primary importance for the aim of this manuscript. We make reference, for simplicity, to the case of discrete-time evolution, where the definition of the adjoint of the evolution operator in tangent space is straightforward§: given  $DL^t(x) : \mathcal{T}_x\mathcal{M} \rightarrow \mathcal{T}_{L^t(x)}\mathcal{M}$  its adjoint operator  $(DL^t(x))^* : \mathcal{T}_{L^t(x)}\mathcal{M} \rightarrow \mathcal{T}_x\mathcal{M}$  represents the time-reversed evolution in tangent space of a reversible dynamical system  $(\mathcal{M}, \mu, L)$ . By definition the operator  $(DL^t(x))^* DL^t(x)$  is a symmetric, positive-definite linear operator on  $\mathcal{T}_x\mathcal{M}$  and, with the same hypotheses of Oseledets' theorem, one can prove that the following limit exists almost everywhere

$$\mathcal{D}(x) = \lim_{t \rightarrow \infty} [(DL^t(x))^* DL^t(x)]^{\frac{1}{2t}}. \quad (7)$$

Moreover, its eigenvectors  $\{d_1(x), \dots, d_N(x)\}$  are a normal basis in  $x$ , while the corresponding eigenvalues  $\{\delta_1(x), \dots, \delta_N(x)\}$  are such that  $\ln \delta_i(x) = \lambda_i(x)$ . Notice that the operator (7) describes the evolution in tangent space forward in time for a duration  $t$  and, then, backward in time for the same lapse of time. Accordingly, the initial and the final tangent spaces coincide, so that the difficulties inherent the extension of the Lyapunov stability analysis to generic trajectories in  $\mathcal{M}$  are removed. Notice also that, while in ergodic systems the eigenvalues of  $\mathcal{D}(x)$  are almost everywhere constant, the corresponding eigenvectors do depend on  $x$ .

## 2.2. Dynamical systems and entropies

For the sake of completeness it is worth discussing shortly the crucial contributions by Ya. B. Pesin to the mathematical theory of LE. In a paper of 1976 [23] he generalized

§Notice *en passant* that this is the general situation one has to deal with when numerical estimates of  $Sp(x)$  and of the corresponding eigenvectors have to be performed. In fact, even continuous-time dynamical systems have to be integrated by transforming them into discrete-time algorithms, whose reliability is primarily related to the conservation of a suitable measure  $\mu$  and its symmetries and to the choice of appropriate integration time steps that guarantee a sufficient sampling all over the manifold  $\mathcal{M}$

the concept of Lyapunov characteristic exponents for a family of mappings that satisfy more general conditions than Lyapunov regularity. In practice, the main result is the identification of invariant manifolds associated to nonzero Lyapunov exponents in non-uniformly hyperbolic systems: this extends to a larger class of dynamical models the study of Lyapunov stability for a measure-preserving dynamics. The main achievements by Pesin are contained in [24]. In that manuscript he proved that the Kolmogorov–Sinai entropy  $K(L)$  of any classical dynamical system  $(\mathcal{M}, \mu, L)$  can be expressed in terms of the LE, through the relation

$$K(L) = A \int_{\mathcal{M}} \sum_i^+ \lambda_i(x) d\mu, \quad (8)$$

where  $A$  is a suitable constant and the sum  $\sum_i^+$  is restricted to the positive  $\lambda_i$ . One of the important consequences of this result for practical applications is that the entropy of a dynamical system (one could say, its *degree of unpredictability*) has a density with respect to the measure  $\mu$  and this density is related to the sum of the positive LE's. For an ergodic system, since the LE are constant almost-everywhere the above expression becomes

$$K(L) = A \sum_i^+ \lambda_i \quad (9)$$

In the same paper [24] Pesin discussed how the mechanism of Oseledets' splitting (i.e. the existence of a filtration of the tangent-space evolution into invariant submanifolds) is related to the hyperbolicity of the dynamical system  $(\mathcal{M}, \mu, L)$ .

### 3. The geometrical structure of tangent space

The existence for invertible dynamics of a coordinate-independent local decomposition of the tangent-space evolution into covariant subspaces (again, the mechanism of Oseledet's splitting) was discussed by Ruelle in a seminal paper of 1979 [5]. Rather than surveying the rigorous mathematical treatment, we have decided to describe here, in a language accessible also to non mathematicians, the main consequences of Ruelle's contribution. For this purpose we shall adopt notations more familiar to an audience of readers interested to applications. In practice, rather than referring to an abstract dynamical system,  $(\mathcal{M}, \mu, L(x))$ , we shall consider explicitly models of physical interest, where  $\mathcal{M}$  is identified with  $\mathbb{R}^N$  and the evolution operator  $L(x)$  is specified by a set of ordinary differential equations

$$\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t) \quad (10)$$

where the point  $x \in \mathcal{M}$  is now represented by the vector  $\mathbf{x}_t = \{x_t^1, x_t^2, \dots, x_t^N\} \in \mathbb{R}^N$  and the continuous dependence on time is made explicit by the subscript  $t \in \mathbb{R}$ , so that the application  $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  is a continuous-time dynamical rule. In this section we limit our analysis to invertible dynamical rules, i.e. we assume that for any given position in the phase-space there exists only one backward trajectory. Since a continuous-time dynamics can be reduced to a discrete-time one with the help of a Poincaré section [26], we make reference only to the map formalism, where the dynamics reads

$$\mathbf{x}_m = \mathbf{f}^{(m)}(\mathbf{x}_0) \quad , \quad \mathbf{x}_{m+n} = \mathbf{f}^{(m)}(\mathbf{x}_n) = \mathbf{f}^{(n)}(\mathbf{x}_m) = \mathbf{f}^{(m)} \mathbf{f}^{(n)}(\mathbf{x}_0) \quad (11)$$

where  $\mathbf{x}_m = \{x_m^1, x_m^2, \dots, x_m^N\} \in \mathbb{R}^N$  and the time variable is an integer index  $m \in \mathbb{Z}$ . Again, the map  $\mathbf{f}^{(m)}$  is assumed to be invertible, i.e. the inverse map  $\mathbf{f}^{(-m)}(\mathbf{x}_m)$  exists and has only one solution.

It is worth pointing out that the assumption of the existence of an invariant measure  $\mu$  is a crucial point that demands much care to be properly formulated¶. In this discrete framework it is convenient to consider the Jacobian matrix at  $\mathbf{x}_n$

$$\mathbf{J}(\mathbf{x}_n) = \frac{\partial \mathbf{f}(\mathbf{x}_n)}{\partial \mathbf{x}_n} \quad , \quad (12)$$

whose elements are assumed to be non-divergent at any  $\mathbf{x}_n$  and defined as follows

$$[\mathbf{J}(\mathbf{x}_n)]_{i,j} = \frac{\partial(x_{n+1}^{(i)})}{\partial x_n^{(j)}} \quad (13)$$

with  $x_n^{(i)} = [\mathbf{x}_n]_i$ . The evolution operator in tangent space  $DL^t(x)$  introduced in Eq. (7) can be represented by the matrix

$$\mathbf{M}_{k,n} = \prod_{i=n}^{k+n-1} \mathbf{J}(\mathbf{x}_i) \quad (14)$$

where  $\mathbf{M}_{k,n} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  and the lapse of time  $t$  has become the integer  $k$ . Note also that with a slight abuse of notation we write  $\mathbf{M}_{k,n}$  instead of  $\mathbf{M}_k(\mathbf{x}_n)$ .

It can be easily shown that  $\mathbf{M}_{k,n}$  satisfies cocycle properties

$$\mathbf{M}_{k+l,n} = \mathbf{M}_{k,n+l} \mathbf{M}_{l,n} \quad , \quad \mathbf{M}_{0,n} = \mathbb{I} \quad \Rightarrow \quad \mathbf{M}_{-k,n} = (\mathbf{M}_{k,n-k})^{-1} \quad (15)$$

The first relation defines the multiplicative property of Jacobian matrices, while the last one defines the inverse, i.e. time-reversed, evolution matrix in the tangent space.

### 3.1. Oseledets' matrix and Oseledets' splitting

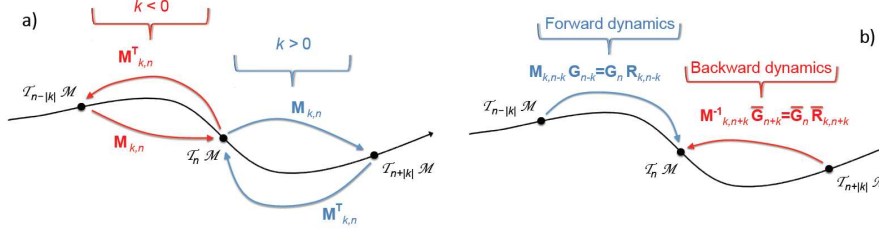
It is worth reformulating Oseledets' theorem by taking advantage of the invertible dynamics and introducing the concepts of forward and backward Oseledets' matrix, denoted by the signs " + " and " - ", respectively. Under the same hypotheses as for Oseledets' theorem, one can prove that the following limits exist

$$\Xi_n^\pm = \lim_{k \rightarrow \pm\infty} \frac{1}{2k} \ln [(\mathbf{M}_{k,n})^T \mathbf{M}_{k,n}] \quad (16)$$

where  $T$  indicates the transpose matrix||. Note that for  $k > 0$  ( $k < 0$ ) the matrix  $(\mathbf{M}_{k,n})^T \mathbf{M}_{k,n}$  evolves forward (backward) a generic tangent space vector from time  $n$  to time  $n + k$ , and then backward (forward) in time up to time  $n$ . Therefore, the forward Oseledets' matrix  $\Xi_n^+$  probes the future dynamics of  $\mathbf{x}_n$ , while the backward

¶In the case of symplectic evolution, ergodicity is usually assumed, in spite of examples where it is known that the ergodic hypothesis does not hold (e.g., the Fermi–Pasta–Ulam model). Moreover, even uniform hyperbolicity or simply hyperbolicity are not guaranteed on rigorous grounds for many models of interest for applications. This notwithstanding, many of the results and tools derived from rigorous ergodic theory reveal effective and very useful, when applied to dynamical models that, presumably, do not fulfill all the hypotheses from which they were derived, as captured by the celebrated chaotic hypothesis[25]. Anyway, one should always keep in mind that a straightforward application of these methods may yield "pathological" outcomes that are just a consequence of singular features emerging from the highly complicated nature of  $\mu$ , provided it does exist.

||We assume that the scalar product is the one associated with the chosen basis. In order to consider a more generic scalar product, one should use the adjoint matrix  $(\mathbf{M}_{k,n})^*$  instead of  $(\mathbf{M}_{k,n})^T$ .



**Figure 1.** (color online) (a) Schematic representation of the action of the symmetric matrices  $\mathbf{M}_{k,n}^T \mathbf{M}_{k,n}$  on tangent space  $\mathcal{T}_n \mathcal{M}$  for finite  $k$ . (b) Forward (+) and backward (-) Gram-Schmidt evolution (23) in tangent space for finite  $k$ . The oriented black line represent the phase space dynamics (10). Blue arrows represent the forward dynamics, red ones the time reversed, backward one.

one  $\Xi_n^-$  probes the past (see Fig. 1).

The forward and the backward Oseledets' matrices share the same  $m \leq N$  distinct eigenvalues  $\lambda_1 > \lambda_2 > \dots > \lambda_m$ , each with multiplicity  $g_1$  ( $\sum_{i=1}^N g_i = N$ ) which coincide with the LEs of the dynamical system (10). The orthogonal eigenvectors  $(\mathbf{d}_n^{(i)})_{\pm}$  (with  $i = 1, \dots, N$ ) of  $\Xi_n^+$  and  $\Xi_n^-$ , on the other hand, do differ and are not invariant under time reversal. As shown in [5], one can construct Oseledets' splitting  $\Omega_n^{(i)}$  ( $i = 1, \dots, m$ ) by simply intersecting the so-called Oseledets' subspaces

$$\Omega_n^{(i)} = (\Gamma_n^{(i)})^+ \cap (\Gamma_n^{(i)})^- \quad (17)$$

where

$$(\Gamma_n^{(i)})^+ = (\mathbf{U}_n^{(i)})^+ \oplus \dots \oplus (\mathbf{U}_n^{(m)})^+ \quad (18)$$

and

$$(\Gamma_n^{(i)})^- = (\mathbf{U}_n^{(1)})^- \oplus \dots \oplus (\mathbf{U}_n^{(i)})^- \quad (19)$$

In both formulae  $(\mathbf{U}_n^{(i)})^{\pm}$  with  $i = 1, \dots, m$  denote the eigenspaces of Oseledets' matrices defined in (16), i.e. they are the (orthogonal) subspaces spanned by the eigenvectors of the forward and backward Oseledets' matrices. Note that, as a consequence of Oseledets' theorem, the subspaces  $(\Gamma_n^{(i)})^{\pm}$  are a filtration of the dynamics (10) or of its time-reversal, that is

$$\lim_{k \rightarrow \pm\infty} \frac{1}{|k|} \ln \frac{\|\mathbf{M}_{k,n} \mathbf{u}\|}{\|\mathbf{u}\|} = \pm \lambda_i \quad \text{for } \mathbf{u} \in (\Gamma_n^{(i)})^{\pm} \setminus (\Gamma_n^{(i\pm 1)})^{\pm} \quad (20)$$

with the nested subspace structure  $\mathbb{R}^N = (\Gamma_n^{(1)})^+ \supset (\Gamma_n^{(2)})^+ \supset \dots \supset (\Gamma_n^{(m)})^+ \supset (\Gamma_n^{(m+1)})^+ \equiv \emptyset$  and  $\mathbb{R}^N = (\Gamma_n^{(m)})^- \supset (\Gamma_n^{(m-1)})^- \supset \dots \supset (\Gamma_n^{(1)})^- \supset (\Gamma_n^{(0)})^- \equiv \emptyset$ . Oseledets' splitting  $\{\Omega_n^{(i)}\}_{i=1,\dots,m}$  is generically non-orthogonal, and determines a *measurable decomposition* of the tangent space which is independent of the chosen norm, covariant with the dynamics and (obviously) invariant under time reversal, that is  $\mathbf{M}_{k,n} \Omega_n^{(i)} = \Omega_{k+n}^{(i)}$ . The Oseledets' decomposition into such *covariant* subspaces exists for any map (11), which is continuous and measurable together with its inverse and whose Jacobian matrix exists and is finite in each element. If the Jacobian matrix is non-degenerate ( $m = N$ ), it can be easily shown that also the spectrum of LEs is non degenerate and all the subspaces  $\Omega_n^{(i)}$  have dimension one. Conversely, if the Jacobian matrix is degenerate, each covariant subspace has a dimension equal to the



degeneracy  $g_i$  of the corresponding LE.

Finally, as a consequence of Eqs. (17,20), it is easy to show that a generic tangent space vector,  $\mathbf{u}_n^{(i)} \in \Omega_n^{(i)}$  for large values of  $k$  grows as

$$\|\mathbf{M}_{k,n}\mathbf{u}_n^{(i)}\| \sim \|\mathbf{u}_n^{(i)}\| \exp^{\lambda_i k} . \quad (21)$$

that is, its exponential asymptotic growth rate is the LE  $\lambda_i$ .

### 3.2. Gram-Schmidt vectors and Lyapunov characteristic exponents

The existence of a well defined limit in Eq. (16) does not allow one to compute neither the spectrum of the Lyapunov characteristic exponents nor the Oseledets' subspaces and splitting by directly diagonalizing the symmetric matrices  $\Xi_n^\pm$ , since they develop exponentially diverging terms leading to numerical overflows. A way out, as noted in Refs. [1, 2], is offered by Eq. (4). The idea is to follow the tangent space evolution of  $p$ -volumes of increasing dimension  $p$ , and to recover the LEs  $\lambda_i$  from Eq. (4). While  $p$ -volumes may be defined by an arbitrary basis of  $p$  independent vectors  $(\tilde{\mathbf{g}}_n^{(1)}, \tilde{\mathbf{g}}_n^{(2)}, \dots, \tilde{\mathbf{g}}_n^{(p)})$ , in order to evolve volumes of dimension  $p > 1$ , one has to take into account the fact that almost every initial vector will converge exponentially towards the largest expanding one,

$$\|\mathbf{M}_{k,n}\tilde{\mathbf{g}}_n^{(i)} - \mathbf{M}_{k,n}\tilde{\mathbf{g}}_n^{(j)}\| \rightarrow 0 \quad \text{for } k \gg 1 \text{ and } \forall i, j, \quad (22)$$

thus quickly making any two vectors  $i$  and  $j$  numerically undistinguishable. To solve this problem it is necessary to periodically orthogonalize the tangent space basis during time evolution. If one considers an orthogonal matrix whose vector columns form an orthonormal basis in tangent space,  $\mathbf{G}_n = (\mathbf{g}_n^{(1)} | \mathbf{g}_n^{(2)} | \dots | \mathbf{g}_n^{(N)})$ , this can be achieved by means of QR decomposition every  $k$  timesteps,

$$\mathbf{M}_{k,n}\mathbf{G}_n = \mathbf{G}_{k+n}\mathbf{R}_{k,n} \quad (23)$$

where  $\mathbf{G}_{k+n}$  is made of orthonormal vectors at  $\mathbf{x}_{k+n} = \mathbf{f}^{(k)}(\mathbf{x}_n) = \mathbf{f}^{(k+n)}(\mathbf{x}_0)$  and the upper diagonal matrix  $\mathbf{R}_{k,n}$  contains the information obtained by the orthonormalization procedure of the *non-orthogonal* matrix  $\tilde{\mathbf{G}}_{n+k} = \mathbf{M}_{k,n}\mathbf{G}_n$ . In particular, its diagonal elements  $[\mathbf{R}_{k,n}]_{i,i} = \gamma_{k,n}^{(i)}$  are the local growth rates over a time  $k$  of the orthogonal vectors  $\mathbf{G}_n$ , that is the growth rates of the axes of inertia of  $p$ -volumes. In numerical applications,  $k$  should not be so large as to produce overflows or problems of numerical accuracy in the QR decomposition of  $\tilde{\mathbf{G}}_{n+k}$ . In systems with an invariant ergodic measure  $\mu$  on  $\mathcal{M}$ , where averages over  $\mu$  can be replaced by time averages, the (ordered) Lyapunov spectrum can be thus computed by taking the time-average of the logarithms of these diagonal terms,

$$\lambda_i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{h=0}^{T-1} \ln \gamma_{k,n+hk}^{(i)} . \quad (24)$$

The off-diagonal nonzero elements  $[\mathbf{R}_{k,n}]_{j,i}$ , on the other hand, are obtained by projecting each vector column  $\tilde{\mathbf{g}}_{n+k}^{(i)}$  onto the subspace spanned by  $\{\tilde{\mathbf{g}}_{n+k}^{(j)}\}$  with  $j < i$ . The orthogonal column vectors of  $\mathbf{G}_m$  evolving under Eq. (23) are often referred to as *forward* Gram-Schmidt (GS) vectors (from the well-known orthonormalization algorithm). In general, they depend on the initial condition  $\mathbf{x}_n$  from which the above procedure has been initialized, but for  $m - n$  sufficiently large they converge to a well defined direction which only depends on  $m$ .

This is the main result obtained by Ershov and Potapov [27]: under the general hypothesis that the considered dynamical system has an invariant ergodic measure, one can show that the orthonormal basis  $\{\mathbf{g}_n^{(i)}\}_{i=1,\dots,N}$  evolved by Eq. (23) converges exponentially to the eigenvectors  $\{(\mathbf{d}_n^{(i)})_-\}_{i=1,\dots,N}$  of the *backward* Oseledets' matrix,

$$\|\mathbf{g}_{n+k}^{(i)} - (\mathbf{d}_{n+k}^{(i)})_-\| \rightarrow 0 \quad \text{for } k \rightarrow \infty. \quad (25)$$

From now on we will always refer to GS vectors as to the “asymptotic” ones\*\*.

An analogous result holds for the *backward* Gram-Schmidt vectors of the time reversed dynamics, which converge to the same index eigenvectors of the *forward* Oseledets' matrix. Obviously, the (reversely ordered) LEs  $\bar{\lambda}_i$  of the time reversed dynamics coincide with the forward ones apart for a sign change

$$\lambda_i = -\bar{\lambda}_i \quad (26)$$

and can be numerically computed by the analogous of Eq. (24). Note that by this convention, the fastest growing direction of the time-reversed dynamics is associated to the  $N$ -th backward GS vector.

It is also important to stress that *the forward and backward Gram-Schmidt vectors are not invariant under time-reversal*. Furthermore, they are norm-dependent objects – as it follows from the norm-dependent nature of GS orthogonalization – with the exception of the first forward and first backward GS vectors (i.e. the ones associated to the largest expansion rate in both time directions).

The fact that the forward GS vectors converge to the eigenvectors of the backward Oseledets' matrix could seem odd at first sight†† but it is easy to realize that both basis in  $\mathcal{T}_n\mathcal{M}$  are only determined by information pertaining to the past dynamics of  $\mathbf{x}_n$ , as shown in Fig. 1. Once again, the same heuristic reasoning applies to backward GS vectors and the eigenvectors of the forward Oseledets' matrix.

#### 4. Covariant Lyapunov vectors

The possibility to characterize and practically compute a norm-independent and time-invariant set of local tangent space vectors associated to the LEs allows one to directly probe the expanding and contracting directions of a given dynamical system. This development has a number of noteworthy applications beyond the area of ergodic theory, ranging from ensemble forecasting to statistical mechanics. It is thus natural to consider the *covariant Lyapunov vectors*  $\mathbf{v}_n^{(i)}$ , which are the unitary vectors spanning the Oseledet's splitting  $\Omega_n^{(i)}$ .

The knowledge of CLVs allows to identify at each point in phase space a vector field that has a natural geometrical interpretation: it is the collection of the tangent space varieties associated to the stability properties of the dynamical system ruled by the LE's. Accordingly, CLVs are objects of primary importance for the study of a dynamical system, since they provide all the information about the local geometrical structure of the tangent space. We now discuss them in detail.

\*\* Ershov and Potapov call the large time limit of the GS vectors a *stationary basis*, a somewhat confusing term since they do vary in time. We therefore prefer to use the term “asymptotic”.

†† In order to avoid this “inversion” between Oseledets eigenvectors and asymptotic GS vectors, some author indeed prefer to name “backward” the GS vectors obtained by moving forward in time, and viceversa. However, since we prefer to emphasize the operational way used to compute them, i.e. by following the forward dynamics, we will stick with our nomenclature.

If  $\dim(\Omega_n^{(i)}) = 1$ , that is, if  $\lambda_i$  is non-degenerate, a single CLV  $\mathbf{v}_n^{(i)}$  associated to the LE  $\lambda_i$  is uniquely defined (up to a phase factor), while for degenerate LEs any nonsingular base formed by  $g_i = \dim(\Omega_n^{(i)})$  covariant vectors  $\mathbf{v}_n^{(i)} \in \Omega_n^{(i)}$  may be arbitrarily considered. In the following, unless otherwise stated, we will unfold eventual degeneracies of the Lyapunov Spectrum, and consider  $N$  LEs  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$  and their corresponding CLVs  $\{\mathbf{v}_n^{(i)}\}_{i=1,2,\dots,N}$ .

In principle, CLVs can be obtained by: i) a “dynamical” approach, which consists in first determining the Lyapunov basis via the GS dynamics (23) forward in time and then following the time-reversed evolution in the corresponding subspaces [9] ; ii) a “static” approach which consists in computing the corresponding Oseledets’s subspaces via (18),(19) and subsequently intersecting them according to Eq. (17) [7]. The intersection of large matrices may be susceptible to problems of numerical accuracy in high-dimensional spaces. Although some variants of the “static” approach, have been recently proposed [10, 11], which improve the performance of such a method, we believe that the former “dynamical” approach is more efficient and we thereby provide a detailed description in following subsection and discuss its stability in Section 5. Anyway, for the sake of completeness, in Section 5, we briefly illustrate also the static algorithms.

#### 4.1. Dynamical algorithm for computing covariant Lyapunov vectors - Formal aspects

The dynamical algorithm sketched in [9] makes use of the following simple but powerful observation. Suppose we have computed the long-time limit of the forward GS vectors (i.e. the asymptotic GS basis) along a certain trajectory. Call these GS vectors  $\mathbf{g}_n^{(i)}$ . Then, the  $j$ -th CLV is contained in the subspace formed by the first  $j$  asymptotic GS vectors. Moreover, a generic vector  $\mathbf{u} \in \mathbb{R}^N$  evolved backward from time  $n$  through the time-reversed dynamics within the corresponding Oseledets’ subspace  $(\Gamma_m^{(j)})_-$  will converge asymptotically to the true CLVs  $\mathbf{v}_m^{(j)}$ . The very fact that CLVs are computed via a stable dynamical rule and not by intersections of large subspaces ensures the numerical stability of this algorithm (see also the next section).

Our first observation follows trivially from the definitions (17-19). It can be expressed as

$$\mathbf{v}_n^{(i)} = \sum_{j=1}^i c_n^{(j,i)} \mathbf{g}_n^{(j)} \quad (27)$$

where the  $c_n^{(j,i)} = \langle \mathbf{g}_n^{(j)} | \mathbf{v}_n^{(i)} \rangle$  with  $j \leq i$  are the CLVs expansion coefficients at time  $n$ . Note that CLVs are defined up to an irrelevant sign (i.e. only their orientation is defined) and that the first CLV coincides with the first vector of the GS basis by construction.

To demonstrate the second point, we first express the CLVs dynamics in a convenient form. Define the matrix whose vector columns are the CLVs,  $\mathbf{V}_n = (\mathbf{v}_n^{(1)} | \mathbf{v}_n^{(2)} | \dots | \mathbf{v}_n^{(N)})$ . Covariant vectors evolve according to

$$\mathbf{M}_{k,n} \mathbf{V}_n = \mathbf{V}_{n+k} \mathbf{D}_{k,n} \quad (28)$$

where the diagonal matrix  $\mathbf{D}_{k,n}$  is composed of the local growth factors  $\gamma_{k,n}^{(i)} = \|\mathbf{M}_{k,n} \mathbf{v}_n^{(i)}\|$ , that is  $[\mathbf{D}_{k,n}]_{i,j} = \delta_{i,j} \gamma_{k,n}^{(i)}$ . For finite  $k$ , the logarithms of these growth factors are called *finite time* Lyapunov exponents (FTLE)

$$\lambda_i^{(k,n)} = \ln \gamma_{k,n}^{(i)} \quad (29)$$

and their time average obviously coincide with the LEs.

For the sake of simplicity, we assume that the spectrum is not degenerate, so that all individual CLVs may be resolved, and we express Eq. (27) in matrix form. We write  $\mathbf{V}_n = \mathbf{G}_n \mathbf{C}_n$ , where the upper triangular matrix  $\mathbf{C}_n$  contains the CLVs expansion coefficients,  $[\mathbf{C}_n]_{j,i} = c_n^{(j,i)}$  for  $j \leq i$ . Note that since the CLVs have unit norm by definition, the expansion coefficients have to be normalized column by column,

$$\sum_{j=1}^i (c_n^{(j,i)})^2 = 1 \quad \forall i. \quad (30)$$

Using Eq. (23) we may rewrite Eq. (28) as

$$\mathbf{G}_{n+k} \mathbf{C}_{n+k} \mathbf{D}_{k,n} = \mathbf{M}_{k,n} \mathbf{G}_n \mathbf{C}_n = \mathbf{G}_{k+n} \mathbf{R}_{k,n} \mathbf{C}_n \quad (31)$$

which implies

$$\mathbf{C}_n = \mathbf{R}_{k,n}^{-1} \mathbf{C}_{n+k} \mathbf{D}_{k,n}. \quad (32)$$

This is our fundamental equation for backward evolution, showing that the CLVs – once expressed in the basis of the GS vectors – can be easily evolved backward by multiplying the inverted upper triangular matrices  $\mathbf{R}_{k,n}^{-1}$ , where  $\mathbf{R}_{k,n}$  is a “by-product” of the forward GS dynamics (23).

Now consider a generic nonsingular upper triangular matrix  $\tilde{\mathbf{C}}_{n+k}$  which, in the GS basis, defines a nonsingular sets of vectors  $\{\tilde{\mathbf{u}}_{n+k}^{(i)}\}_{i=1,\dots,N}$  projected in the Oseledets’ subspaces, such that  $\tilde{\mathbf{u}}_{n+k}^{(i)} \in (\Gamma_{n+k}^{(i)})_-$ . We compare its backward evolution with that of true CLVs,  $\mathbf{u}_{n+k}^{(i)}$  (which in the GS vectors basis we write in the matrix form  $\mathbf{C}_{n+k}$ ). Evolving both sets of vectors backward according to Eq. (32), we have

$$\tilde{\mathbf{C}}_n = \mathbf{R}_{k,n}^{-1} \tilde{\mathbf{C}}_{n+k} \tilde{\mathbf{D}}_{k,n} \quad (33)$$

$$\mathbf{C}_n = \mathbf{R}_{k,n}^{-1} \mathbf{C}_{n+k} \mathbf{D}_{k,n} \quad (34)$$

which implies

$$\tilde{\mathbf{C}}_{n+k} \tilde{\mathbf{D}}_{k,n} \tilde{\mathbf{C}}_n^{-1} = \mathbf{C}_{n+k} \mathbf{D}_{k,n} \mathbf{C}_n^{-1} \quad (35)$$

and thus

$$\tilde{\mathbf{C}}_n^{-1} \mathbf{C}_n = \tilde{\mathbf{D}}_{k,n}^{-1} \tilde{\mathbf{C}}_{n+k}^{-1} \mathbf{C}_{n+k} \mathbf{D}_{k,n}. \quad (36)$$

It is easy to verify that

$$[\tilde{\mathbf{C}}_n^{-1} \mathbf{C}_n]_{i,j} = \frac{\gamma_{k,n}^{(j)}}{\tilde{\gamma}_{k,n}^{(i)}} [\tilde{\mathbf{C}}_{n+k}^{-1} \mathbf{C}_{n+k}]_{i,j} \quad (37)$$

where both sides of Eq. (37) are upper triangular matrices. Thus, the r.h.s. matrix-elements are zero for  $i > j$ . On the other hand, for  $i < j$  we have

$$\frac{\gamma_{k,n}^{(j)}}{\tilde{\gamma}_{k,n}^{(i)}} \rightarrow 0 \quad \text{for } k \rightarrow \infty \text{ and } i < j \quad (38)$$

due to the nested structure of the Oseledets’ subspaces. For  $k \gg 1$  one has (up to an irrelevant sign)

$$[\tilde{\mathbf{C}}_n^{-1} \mathbf{C}_n]_{i,j} \approx \delta_{i,j} [\tilde{\mathbf{C}}_{n+k}^{-1} \mathbf{C}_{n+k}]_{j,j} \equiv \delta_{i,j} q_j \quad (39)$$

where  $q_j$  are arbitrary factors which solely depend on the angles between the initial vectors at time  $n + k$ . We finally have, for  $k \gg 1$ ,

$$[\tilde{\mathbf{C}}_n]_{i,j} \approx [\mathbf{C}_n]_{i,j} q_j^{-1} \quad (40)$$

that is, under the backward evolution (28) almost any nonsingular vector set (once it belongs to the Oseledets' subspace) converges backward in time towards the true CLV, apart from a trivial scale factor. This proves our second point.

As a final practical remark, we want to point out that numerical estimates of LE and CLV are always performed over long but obviously finite time lapses, whereas all theorems hold in the infinite time limit. In the next Section, we show that the convergence rate is typically exponential, though the rate may be rather small for large system sizes. Accordingly, we do not expect any practical problem.

#### 4.2. Dynamical algorithm for computing covariant Lyapunov vectors - Numerical implementation

We now discuss the numerical aspects of the dynamical algorithm. The first part closely follows the standard algorithm introduced by Benettin *et al.* to compute LEs [2], while in the second part, backward evolution is implemented via Eq. (28).

- (i) *Forward transient* – Given a generic initial condition  $\mathbf{x}_0$  in phase space and a generic set of  $m \leq N$  orthogonal tangent-space vectors  $\{\mathbf{g}_0^{(i)}\}_{i=1,\dots,m}$ , we first evolve the phase and the tangent space dynamics via Eq. (10) and Eq. (23), respectively. This transient should last a number  $n$  of timesteps sufficient for the phase-space trajectory to converge to the ergodic attractor and for the orthogonal vectors to converge to the asymptotic GS vectors.
- (ii) *Forward dynamics* – Once a GS basis is reached, we proceed further evolving the reference trajectory  $\mathbf{x}_n$  and GS vectors by  $k$  timesteps at once, recording in memory both the local GS vectors  $\mathbf{G}_{n+hk}$  and the upper triangular matrices  $\mathbf{R}_{k,n+(h-1)k}$ , for  $h = 1, 2, \dots, t + t_0$ .
- (iii) *Backward transient* – A generic non-singular upper triangular matrix  $\tilde{\mathbf{C}}_{n+(t+t_0)k}$  is generated and evolved backward via Eq. (28) for  $t_0$  steps (each moving backward by  $k$  timesteps). The backward transient length  $kt_0$  should be sufficient to converge the tangent space initial conditions close enough to the true CLVs expansion coefficients in  $\mathbf{x}_{n+tk}$ . Columns of  $\tilde{\mathbf{C}}_{n+(t+t_0)k}$  are kept normalized via the condition (30).
- (iv) *Backward dynamics* – Finally, the CLVs expansion coefficients may be further evolved backward along the trajectory (always being kept normalized) to sample the geometrical structure of the ergodic attractor at points  $\mathbf{x}_{n+hk}$ , with  $h = 0, 1, \dots, t$ . CLVs expressed in the phase space coordinates reference can be obtained by the GS vectors  $\mathbf{G}_{n+hk}$  via Eq. (27).

Several comments are in order:

- *Sampling frequency* – Note that  $k$  dictates the sampling frequency of CLVs over the phase space, reference trajectory  $\mathbf{x}_{n+hk}$ . It can be adjusted as needed, but it should be small enough to avoid numerical overflows and too large numerical errors in the  $\mathbf{R}_{k,m}$  and  $\mathbf{C}_m$  matrices due to exponential growth and contractions. The choice of  $k$  thus obviously depends on both the stability properties of the chosen dynamical system and on the required sampling frequency. In principle,

two different values can be chosen, one ( $k_{tr}$ ) for the forward and backward transients, and a second  $k_{dyn} < k_{tr}$  for the forward and backward dynamics.

- *Scalability* – The algorithm is fully scalable, and if one just needs to compute the first  $m \leq N$  CLVs, that is the  $m$  most expanding vectors, no other vectors with  $i > m$  are needed. Indeed, in order to compute the first  $m$  CLVs one only needs to consider the  $N \times m$  reduced GS matrices  $\mathbf{G}_n$  and the corresponding  $m \times m$  reduced upper triangular matrices  $\mathbf{R}_n$ . If the dynamics is explicitly invertible, the same reasoning could be applied to the computation of the last  $m$  CLVs, i.e. the  $m$  most contracting vectors, by simply writing the dynamical algorithm for the time-reversed dynamics (provided that the right trajectory has been generated forward in time).
- *Degenerate Lyapunov spectra* – If the Lyapunov spectrum is degenerate, and some covariant subspaces  $\Omega_m^{(i)}$  have a dimension  $g_i$  larger than one, the individual vectors spanning such a degenerate subspace have no physical meaning. In this case, the above algorithm will simply return some arbitrary basis of  $\Omega_m^{(i)}$  which may depend on the initial choice of  $\tilde{\mathbf{C}}_{n+(t+t_0)k}$ . For a more detailed discussion of the structure of the degenerate CLVs associated to zero LE in a quasi-one dimensional Hamiltonian systems of hard disks see for instance Ref. [18].
- *Memory issues* – All existing algorithms need to store a large amount of information. Our dynamical algorithm requires to store the  $\mathbf{R}_{k,m}$  and  $\mathbf{G}_m$  matrices at periodic sampling points during the forward dynamics, in order to be reused during the backward evolution. The matrices  $\mathbf{R}_{k,m}$  needed to run the backward dynamics involve  $m(m+1)/2$  floating point numbers at each of the  $t + t_0$  sampling points, while the  $\mathbf{G}_m$  matrices involve  $mN$  floating-point numbers for  $t$  sampling points. In terms of floating-point numbers, the total memory requirement is thus

$$M_{tot} = t m \left[ N + \frac{m+1}{2} \right] + t_0 \frac{m(m+1)}{2}. \quad (41)$$

This burden can be sensibly reduced if one is not interested in the spatial structure of CLVs but just in their relative angles, since the  $\mathbf{G}_m$  matrices are not needed (the relative angles can be computed directly with reference to the GS vectors basis). In this case, the memory requirement is just

$$M'_{tot} = (t + t_0) \frac{m(m+1)}{2}. \quad (42)$$

These estimates could nevertheless turn into a large amount of memory for long sampling times, large dynamical systems and/or when many CLVs are required, quickly exhausting the available fast-access memory. Since accessing disk storage every  $k$  timesteps could be rather time-consuming (and disk memory in itself is a limited resource), we suggest an alternative strategy. Once determined the maximum amount of data  $M_b$  that can be stored in the fast-access memory, divide the total forward integration time  $k(t + t_0)$  into  $n_b$  blocks of length  $h_b k$  such that  $M_{tot}/n_b \leq M_b$ . After running the usual forward transient in order to ensure a proper convergence of the GS vectors at time 0, perform a first forward-dynamics run, without storing the  $\mathbf{R}_{k,m}$  and  $\mathbf{G}_m$  matrices every  $k$  timesteps as usual, but just saving the current phase space configuration and the GS vectors every  $h_b k$  timesteps at times  $nh_b k$ , with  $n = 0, 1, \dots, n_b - 1$ , that is, at the *beginning* of each block (this means one does not have to run forward in time

the last block). These data can be typically stored on disk, as it will not be accessed too frequently. Once the first forward run is concluded, perform a series of forward and backward iterations block-by-block, *starting from the last block*. Thus, recover the phase space trajectory and GS vectors at time  $(n_b - 1)h_b k$ , evolve them forward up to time  $n_b h_b k$ , now storing in the fast access memory the  $\mathbf{R}_{k,m}$  and  $\mathbf{G}_m$  matrices every  $k$  timesteps, and subsequently evolve backward a randomly generated nonsingular initial condition  $\tilde{\mathbf{C}}_{n_b h_b k}$ . Next, pass to the  $(n_b - 1)$ -th block, accessing the data stored at time  $(n_b - 2)h_b k$ , performing the forward evolution up to time  $(n_b - 1)h_b k$ . When coming back, however, restart from the coefficient computed in the previous block,  $\tilde{\mathbf{C}}_{(n_b - 1)h_b k}$ , so that the convergence to the true CLVs is not lost. This procedure is finally repeated block by block up to the first one, greatly reducing the amount of memory, at the only expense of performing twice the same forward and backward dynamics, thus doubling the algorithm computational time.

- *Computational time* – We give order-of-magnitude estimates for the above dynamical algorithm in terms of elementary arithmetic operations (basically multiplication/division floating point operations). Suppose we have a dynamical system with  $N$  degrees of freedom and we are interested in computing the first  $m$  CLVs.

Forward dynamical evolution requires to run both phase and tangent space dynamics for  $k$  timesteps and then to perform a single QR decomposition (or Gram-Schmidt orthonormalization). Special dynamical systems such as the ones with strictly local or globally coupled interactions (in the following: easy dynamics) require only  $\mathcal{O}(N)$  operations for a single step of phase space dynamics and  $\mathcal{O}(mN)$  for the tangent space evolution. Unfortunately, generic dynamical systems with long-range interactions (in the following: hard dynamics) typically require  $\mathcal{O}(N^2)$  and  $\mathcal{O}(mN^2)$  operations for phase-space and tangent-space dynamics $\ddagger\ddagger$  respectively (the latter with a prefactor  $\leq 1$ ). This leads to  $\mathcal{O}(mkN)$  (for easy dynamics) or  $\mathcal{O}(mkN^2)$  (for hard dynamics) operations to be performed between consecutive QR decompositions, which is an  $\mathcal{O}(m^2N)$  algorithm in itself. In particular, the stabilized Gram-Schmidt algorithm requires  $\sim 2m^2N$  operations [28].

Backward evolution of the CLVs coefficients via Eq. (32) does not require any explicit matrix inversion, but can be efficiently computed by back-substitution algorithms in  $\sim m^3/3$  operations, while the normalization (30) requires  $\mathcal{O}(m^2)$  operations. Finally,  $\sim m^2N/2$  operations are needed in order to express the CLVs in the phase space coordinate basis via Eq. (27).

Since the convergence of GS vectors and CLVs is exponential (see Section 5 below), and in typical applications in systems with many degrees of freedom (where computational time may become an issue) one is interested to sample the geometrical structure of tangent space over the entire ergodic attractor, one usually has  $t \gg 1$  and  $k_{\text{dyn}} \ll m$ . For easy dynamics (which nevertheless covers the large majority of applications) this implies that the forward evolution is dominated by the QR decomposition, while the backward dynamics by the back substitution, for a total computational time

$$T_{\text{tot}} \approx t m^2 \left( \frac{m}{3} + \frac{N}{2} + 2N \right) = t m^2 \left( \frac{m}{3} + \frac{5}{2}N \right). \quad (43)$$

$\ddagger\ddagger$ We assume the usage of standard matrix multiplication algorithms.

Note that even when the full set of CLVs is computed,  $m = N$ , the backward evolution is approximately 2.4 times faster than the forward one, or 6 times if one is not interested in expressing the CLVs in the phase space coordinate basis. The situation is slightly less favorable for hard dynamics, which can be dominated by the forward evolution in tangent space, requiring full matrix multiplications every time step, for a total time  $T_{\text{tot}} \sim t k m N^2$ .

- *Transient length* – As mentioned in the previous subsection, the convergence rates in both the forward and backward transient rates are exponential and related to the difference between consecutive Lyapunov exponents (see also Section 5 below). For instance, the convergence to the  $i$ -th CLV ( $i > 1$ ) depends on the difference  $\lambda_i - \lambda_{i-1}$ . Dynamical systems with many degrees of freedom are typically characterized in the  $N \rightarrow \infty$ , by a limit spectrum  $\lambda(i/N)$  that is piecewise continuous [29] (possibly after removing those exponents connected to collective and sub extensive modes [15, 30]). This implies that the difference between consecutive exponents scales to zero as  $1/N$ , so that it is advisable, when performing a finite size analysis of such systems, to scale the transient time with the number of degrees of freedom,  $t_0 \rightarrow t_o N$ .

#### 4.3. Non-invertible dynamics

We finally discuss the extension of the concept of CLVs and of the dynamical algorithm to non-invertible dynamics. Mathematically, the construction of Oseledets' splitting requires the uniquely determined future and past images of any phase space point  $\mathbf{x}_n$  on the attractor. Although it is impossible to meet such a requirement in a non-invertible dynamical system (such as, e.g. a chain of logistic maps), the difficulty may be circumvented by identifying the past images as *those ones that have been actually visited during the forward evolution* leading to  $\mathbf{x}_n$ . In this way, we can “artificially” restore the uniqueness of past images, and be free to apply all the machinery described above for invertible dynamics.

There is of course a price to pay: some points  $\mathbf{x}_n$  on the ergodic attractor can be reached by following different past trajectories. This means that CLVs are not uniquely defined, but may depend on the past trajectory followed to reach a given point. There is no longer a unique CLV in each point in phase space. This is not really an obstacle whenever one, instead of being interested in reconstructing the local tangent-space structure, aims at determining general statistical properties of the CLVs.

### 5. Convergence towards the covariant Lyapunov vectors

We now analyze more closely the convergence of the dynamical algorithm towards the CLVs, making use of fluctuations of finite time Lyapunov exponents. Let us assume that the forward evolution has been carried out for a long-enough time to ensure a proper convergence to the maximally expanding subspaces (i.e. the GS vectors) along a given trajectory  $\mathbf{f}^{(n)}(\mathbf{x}_0)$  for  $n = 0, 1, \dots, t + t_0$ , and that the CLVs  $\mathbf{v}_n^{(i)}$  are also known at each point  $x_n$  for  $n = 0, 1, \dots, t$ . If  $t_0 \gg 1$  we can assume these CLVs to be perfectly converged.

Now repeat the backward procedure starting for some randomly chosen vector at time



$$n + k \ll t + t_0 - n,$$

$$\mathbf{u}_{n+k}^{(i)} = \sum_{j=1}^i \tilde{c}_{n+k}^{(j,i)} \mathbf{g}_{n+k}^{(j)} \quad (44)$$

and evolve backward up to time  $n$ . This procedure, allows to study the convergence towards  $\mathbf{v}_n^{(i)}$  of the “approximate” CLV  $\mathbf{u}_n^{(i)}(k)$  as a function of the finite backward evolution time  $k$  on which the vector depends. Let us define  $\delta u_n^{(i)}(k) = |\mathbf{u}_n^{(i)}(k) - \mathbf{v}_n^{(i)}|$  which, being both vectors with unit norm, is just a function of their relative angle  $\theta_n^{(i)}(k)$ ,

$$\delta u_n^{(i)}(k) = \sqrt{2} \sqrt{1 - \cos \theta_n^{(i)}(k)} \quad (45)$$

(note that  $\delta u_n^{(i)}(k) \sim \theta_n^{(i)}(k)$  for small angles).

Since we are interested in the average convergence rate towards the CLVs, we consider the generalized distance

$$\Delta(k, q) = \langle [\delta u_n^{(i)}(k)]^q \rangle^{1/q} \quad (46)$$

where  $\langle \cdot \rangle$  denotes an average over both the ending points  $\mathbf{x}_n$  and the different initial conditions  $\tilde{c}_{n+k}^{(j,i)}$  in tangent space. Moreover, in order to lighten the notations, we have dropped the dependence on the index  $i$  (that is kept only when strictly necessary).

Let us now consider a chain of Hénon maps [31],

$$x_{n+1}(\ell) = a - [x_n(\ell) + \varepsilon(x_n(\ell - 1) - 2x_n(\ell) + x_n(\ell + 1))]^2 + bx_{n-1}(\ell), \quad (47)$$

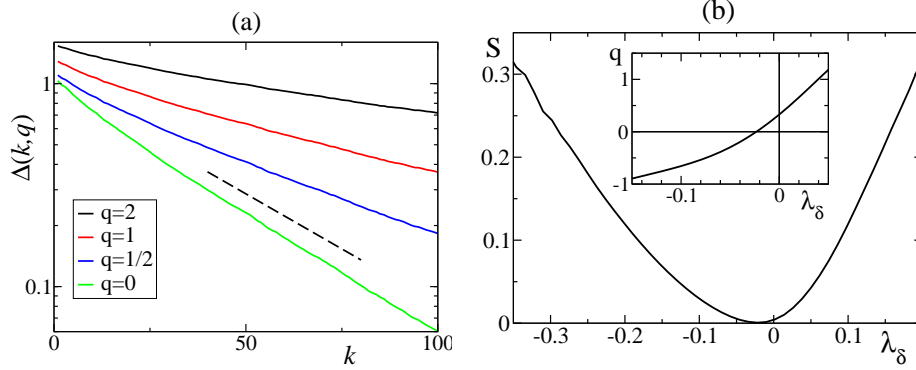
where  $i = 1, \dots, N$  is a lattice index and periodic boundary conditions are assumed. We have studied a set of  $N = 5$  maps for  $a = 1.4$ ,  $b = 0.3$ , and  $\varepsilon = 0.025$ . For such parameter values there are 5 positive Lyapunov exponents and 5 (quite large) negative ones. We focus on the convergence towards the second CLV following the procedure sketched above (the first one is automatically generated during the forward iteration).

In Fig. 2a we plot the behaviour of the generalized distance (46) as a function of the backward convergence time for four different  $q$  values. We see that the different  $q$ -distances all converge to 0 exponentially, thus confirming the quick convergence of the method. However, different  $q$ -distances converge with different exponential rates. In order to describe the process, it is reasonable to conjecture that the evolution towards the  $i$ -th CLV, as given by Eqs. (37-38), is ultimately controlled by the largest ratio between the various subspaces growth factors  $\gamma_{k,n}^{(i)}$ . This is tantamount to assuming that all other off-diagonal terms have already converged to zero and at some time  $m < n + k$  the vector  $\mathbf{u}_m^{(i)}(k)$  lies in the plane defined by the  $(i - 1)$ -th and the  $i$ -th CLVs.

In this plane, the dynamics is controlled by the fluctuations of the finite time difference of the two FTLE,  $\lambda_\delta = \lambda_i(k) - \lambda_{i-1}(k)$  (we implicitly assume that the spectrum is not degenerate in  $i$ ), which for large  $k$  converges to the difference between the two corresponding Lyapunov exponents,  $\Delta\lambda = \lambda_i - \lambda_{i-1} < 0$ . Therefore, for large times,  $\mathbf{u}_m$  tend to rotate towards the  $i$ -th CLV and their relative angle  $\theta_m(k)$  approaches zero for a large majority of initial conditions.

We now consider the restricted generalized distance

$$\Theta^q(k, q) \simeq \langle \theta_m(k)^q \rangle \quad (48)$$



**Figure 2.** (color online) Convergence of the second CLV in a chain of 5 coupled Hénon maps (periodic boundary conditions) as defined in Eq. (47) for  $a = 1.4$ ,  $b = 0.3$  and  $\varepsilon = 0.025$ . The uncertainty  $\Delta(k, q)$ , defined in Eq. (46), is plotted in panel a for different  $q$ -values. The dashed black line marks exponential decay with a rate  $\lambda_2 - \lambda_1$ . In panel b, the corresponding large deviation function is plotted (by following the procedure discussed in the text). Finally, the inset contains the corresponding curve  $q(\lambda_\delta)$ .

where the average has been restricted to small angles only. For small angles we may also write

$$\theta_m(k) \simeq \theta_0 e^{\lambda_\delta k}. \quad (49)$$

Accordingly, the fluctuations of  $\lambda_\delta$  can be described by a large-deviation function [32]  $S(\lambda_\delta)$ ,

$$P(\lambda_\delta, k) \approx e^{-S(\lambda_\delta)k} \quad (50)$$

with  $S(\Delta\lambda) = 0$  being the minimum of  $S$ . Thereby,

$$\Theta^q(k, q) = \int_{-\infty}^0 d\lambda_\delta P(\lambda_\delta, k) \theta_m(k)^q = \int_{-\infty}^0 d\lambda_\delta e^{k[\lambda_\delta q - S(\lambda_\delta)]} \equiv e^{L(q)k} \quad (51)$$

where, the upper bound to the integration interval is needed to avoid the unphysical divergence of the angle. As long as the maximum of the exponential integrated in Eq. (51) is reached for a negative  $\bar{\lambda}_\delta$  (which is itself a function of  $q$ ), that is, if the maximum lies inside the integration interval,  $L(q)/q$  coincides with the generalized Lyapunov exponent (difference)  $\Lambda_q$  [33], and

$$L(q) = \Lambda_q \equiv \bar{\lambda}_\delta q - S(\bar{\lambda}_\delta) \quad , \quad q = S'(\bar{\lambda}_\delta). \quad (52)$$

It is easy to check that for  $q \rightarrow 0$ ,  $\Lambda_q/q$  converges to the Lyapunov exponents difference  $\Delta\lambda$  (which is the extremal point where  $S(\lambda_\delta = \Delta\lambda) = 0$ ).

On the other hand, if  $S(\lambda_\delta)$  extends to the positive semi-axis, there always exists a critical

$$q_c = S'(0) \quad (53)$$

above which, the maximum is attained for  $\lambda_\delta = 0$  and

$$L(q) = -S(0). \quad (54)$$

To resume, we have that  $L(q)$  is non positive and monotonically decreasing for  $0 \leq q \leq q_c$ , from  $L(0) = 0$  to  $L(q_c) = -S(0)$ . For  $q > q_c$ , finally, we have

$$L(q) = -S(0).$$

However, this is not yet the end of the story, since so far we have restricted our generalized distance to small angles. In fact, there also exists a family of special initial conditions  $\theta_0$  that are still close to the  $(i-1)$ -th CLV – i.e. the “wrong” direction – at time  $k$ . They are those for which  $\theta_0$  is extremely close to  $\pi/2$ . The fraction of such trajectories is  $\exp(\lambda_\delta k)$ , if  $\lambda_\delta < 0$ , and is of order 1 otherwise. Altogether, their contribution  $\bar{\Theta}^q(k, q)$  to the total generalized distance is

$$\bar{\Theta}^q(k, q) = \int_{-\infty}^0 d\lambda_\delta e^{k[\lambda_\delta - S(\lambda_\delta)]} = e^{L(1)k}. \quad (55)$$

We are now in a position to evaluate the two contributions combined,

$$\Delta_q(k) = [\Theta^q(k, q) + \bar{\Theta}^q(k, q)]^{1/q}. \quad (56)$$

If  $q_c < 1$  then  $L(1) = L(q_c)$  and the contribution from  $\theta_0 \simeq \pi/2$  never prevails.

If  $q_c > 1$ , on the other hand, the large angles contributions prevails for  $q \geq 1$ . Note also that if the large deviations are restricted to negative values,  $q_c$  does not exist (being formally infinite) and this latter case applies. To sum up, the generalized distance decays exponentially as

$$\Delta(k, q) \sim \begin{cases} e^{k L(1)/q} & \text{for } q \geq \min(q_c, 1) \\ e^{k \Lambda_q/q} & \text{for } q < \min(q_c, 1) \end{cases} \quad (57)$$

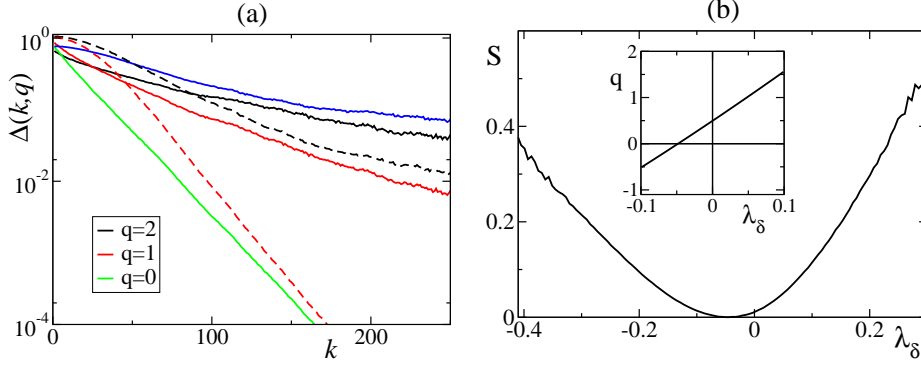
where  $q_c$  is given by Eq. (53). Furthermore, for  $q \rightarrow 0$ , the exponential convergence rate to the  $i$ -th CLV ( $i > 1$ ) is given by the Lyapunov exponent difference  $\Delta\lambda_i = \lambda_i - \lambda_{i-1}$ , while it decreases monotonically as  $q$  is increased.

In Fig. 2b we plot the large deviation function for  $\lambda_\delta = \lambda_2 - \lambda_1$  for the above mentioned chain of Hénon maps. In order to minimize finite-size effects,  $S(\lambda_\delta)$  has been estimated by comparing the histograms for two different time lengths  $k_1 < k_2$ ,

$$S(\lambda_\delta) = \frac{\log P(\lambda_\delta, k_2) - \log P(\lambda_\delta, k_1)}{k_2 - k_1}. \quad (58)$$

The result for  $k_1 = 30$  and  $k_2 = 40$  is plotted in Fig. 2b. In the inset we plot the derivative of  $S$ , which coincides with  $q$ . The crossing of  $q(\lambda_\delta)$  with the horizontal axis identifies the minimum of  $S$ , i.e. the long time average of  $\lambda_\delta$  that turns out to be rather consistent with the direct estimate of  $\Delta\lambda = \lambda_2 - \lambda_1$  ( $-0.023$  instead of  $-0.025$ ). The crossing with the vertical axis identifies  $q_c$ , which, in this case is  $q_c = 0.3$ , so that we are in the case  $q_c < 1$ .

We are now in the position to compare the direct results plotted in Fig. 2a with the prediction of the multifractal analysis. First of all we note that the exponential decay for  $q = 0$  is in perfect agreement with the Lyapunov exponents difference  $\Delta\lambda$  (see the dashed line). As for the other  $q$ -distance, since they are all larger than  $q_c$  their exponential decay rates are expected to be equal to  $S(0)/q$ . By computing the exponential rate for the largest times, we find that they are consistent with the predicted  $1/q$  behavior, although with an estimated prefactor value of  $S(0) \approx 0.008$ . On the other hand a direct estimate from Fig. 2b yields  $S(0) = 0.004$ : the prefactor difference is relatively large, but definitely compatible with the amplitude of the finite-size corrections. Notice that since the minimum of  $S(\lambda_\delta)$  occurs for a small  $\lambda_\delta$  value,  $S(0)$  is quadratically small and this implies that a direct estimate is rather problematic. To test our prediction in a different system, we have also studied a Hamiltonian



**Figure 3.** (color online) Convergence of the 3rd and 9th CLVs in a chain of 5 FPU oscillators and energy density  $e = 5$ . (a) Uncertainty  $\Delta(k, q)$  for different  $q$ -values and both the 3rd (solid lines) and 9th (dashed line) vectors. (b) Large deviation function  $S$  for the 3rd CLV. In the inset the corresponding curve  $q(\lambda_\delta)$  is shown. Additional technical details are mentioned in the text.

model, namely a chain of Fermi-Pasta-Ulam (FPU) oscillators [34], again with periodic boundary conditions,

$$\ddot{q}_i = F(q_{i+1} - q_i) - F(q_i - q_{i-1}) \quad (59)$$

where  $F(x) = x - x^3$ . In this model the energy  $H = \sum \dot{q}_i^2/2 + (q_{i+1} - q_i)^2/2 + (q_{i+1} - q_i)^4/4$  is conserved. We have studied a chain with  $N = 5$  oscillators (and thus 10 degrees of freedom) and energy density  $e = H/N = 5$ . The convergence of the 3rd and 9th covariant vectors are reported in Fig. 3a (see solid and dashed lines respectively). The 0-distance convergence of the two vectors is controlled by the differences  $\lambda_3 - \lambda_2$  and  $\lambda_9 - \lambda_8$ , respectively. Given the symmetry of the Lyapunov spectrum ( $\lambda_i = \lambda_{N+1-i}$ ), the two vectors are expected to converge asymptotically in the same way, and this is another way of checking the correctness of our approach. In fact, we see that the three pairs of curves corresponding to different  $q$ -values become eventually parallel. The initial difference is due to the fact that the 9th vector actually evolves in a higher dimensional space, where additional (and faster) relaxation processes are present.

Furthermore, we have studied the large deviation function for the difference between the 3rd and 2nd Lyapunov exponent (by comparing finite-time Lyapunov exponents for a time  $t_1 = 10$  and  $t_2 = 20$ ). The results are plotted in Fig. 3b. There we see that the minimum of  $S(\lambda_\delta)$  occurs for  $\lambda_\delta = -0.05$  that is quite close to the asymptotic value  $\Delta\lambda = -0.053$ . Moreover, we have  $q_c \approx 0.5$ , i.e. this case is borderline between the two above mentioned classes. Finally, the “asymptotic” exponential rates for the curves with  $q = 1$  and  $q = 2$  in Fig. 3a are  $-0.066$  and  $0.014$ , hinting at  $S(\lambda_\delta) \approx 0.013$ , a value close to the direct observation  $0.011$  (from the inset in Fig. 3b), thus confirming our theoretical argument.

To summarize, we have shown that our algorithm converges exponentially to the true CLVs. However, when decay rates are computed via finite time ensemble averages, only the 0-norm decay rate coincides with the difference between two consecutive Lyapunov exponents. Higher norm decay with smaller exponential rates due to fluctuations of the FTLEs. The norm-dependence can be explained in the context of large deviation theory.

## 6. Static algorithms for covariant Lyapunov vectors

As anticipated, there exist various algorithms for the computation of CLVs. An alternative to the dynamical approach described in Sec. 4 is represented by “static” algorithms, which do not make use of the intrinsic stability when the backward evolution is restricted to suitable subspaces, but rather, determine the CLVs as linear combinations of either forward or backward GS vectors. More precisely, at each point  $\mathbf{x}$  along a given trajectory one has (for the sake of simplicity, in this section we omit the time index)

$$\mathbf{v}^{(i)} = \sum_{j=1}^i \langle \mathbf{g}_+^{(j)} | \mathbf{v}^{(i)} \rangle \mathbf{g}_+^{(j)} = \sum_{j=i}^N \langle \mathbf{g}_-^{(j)} | \mathbf{v}^{(i)} \rangle \mathbf{g}_-^{(j)} \quad (60)$$

where  $\mathbf{g}_+^{(j)}$  and  $\mathbf{g}_-^{(j)}$  are, respectively, the forward and backward asymptotic GS vectors.

### 6.1. Wolfe-Samelson algorithm

A first static algorithm was introduced by Wolfe and Samelson in Ref. [10]. From Eq. (60), by using of the identity relation

$$\sum_{k=1}^N \langle \mathbf{g}_+^{(i)} | \mathbf{g}_-^{(k)} \rangle \langle \mathbf{g}_-^{(k)} | \mathbf{g}_+^{(j)} \rangle = \delta_{i,j} \quad (61)$$

and with the help of simple algebraic manipulations, one obtains (see [10] for more details)

$$\sum_{j=1}^h \sum_{k=i}^{h-1} \langle \mathbf{g}_+^{(i)} | \mathbf{g}_-^{(k)} \rangle \langle \mathbf{g}_-^{(k)} | \mathbf{g}_+^{(j)} \rangle c_+^{(j,h)} = 0 \quad i \leq h \quad (62)$$

where  $c_+^{(j,h)} = \langle \mathbf{g}_+^{(j)} | \mathbf{v}^{(h)} \rangle$  is the CLV coefficient expansion on the forward GS basis. Eq. (61) can be recast in a matrix form, introducing the  $h \times h$  square matrix

$$[\mathbf{Q}^{(h)}]_{ij} = \sum_{k=i}^{h-1} \langle \mathbf{g}_+^{(i)} | \mathbf{g}_-^{(k)} \rangle \langle \mathbf{g}_-^{(k)} | \mathbf{g}_+^{(j)} \rangle \quad i, j \leq h \quad (63)$$

and the vector

$$[\mathbf{y}^{(h)}]_k = \langle \mathbf{g}_+^{(k)} | \mathbf{v}^{(h)} \rangle \quad k \leq h, \quad (64)$$

namely,

$$\mathbf{Q}^{(h)} \mathbf{y}^{(h)} = 0. \quad (65)$$

Eqs. (62) and (65) hold for  $h > 1$  (the first CLV trivially coincides with the first forward GS vector) and allow obtaining the expansion coefficients of the  $h$ -th CLV as the kernel of a matrix computed from the first  $h$  forward and  $h-1$  backward asymptotic GS vectors (the latter being the stable ones for the time-reversed dynamics).

### 6.2. Kuptsov-Parlitz algorithm

More recently, Kuptsov and Parlitz [11] have introduced a similar procedure, which makes use of LU factorization.

By using the matrix notation introduced in Section 4, Eq. (60) can be rewritten as

$$\mathbf{V} = \mathbf{G}_+ \mathbf{C}_+ = \mathbf{G}_- \mathbf{C}_- \quad (66)$$

with the plus and minus indices referring to forward and backward dynamics, respectively. Since the forward matrix  $\mathbf{C}_+$  is upper triangular, while the backward one  $\mathbf{C}_-$  is lower triangular, Eq. (66) can be easily recast as an LU factorization [28]

$$\mathbf{P} \mathbf{C}_+ = \mathbf{C}_- \quad (67)$$

where

$$\mathbf{P} = \mathbf{G}_-^T \mathbf{G}_+ \quad (68)$$

Once again, if we are only interested in the  $j$ -th CLV, that is, in the  $j$ -th column of the matrix  $\mathbf{C}_+$ , only the  $(h-1) \times h$  upper left corner of matrix  $\mathbf{P}$  is needed. Since  $\mathbf{C}_-$  is lower triangular and the first  $j$  entries of its  $j$ -th column are all zeros, we are left with the following system of  $(j-1)$  linear homogeneous equations in  $j$  variables

$$\sum_{i=1}^j [\mathbf{P}]_{k,i} [\mathbf{C}_+]_{i,j} = 0 \quad k = 1, 2, \dots, j-1 \quad (69)$$

which defines the vector  $j$  up to a rescaling factor. Not surprisingly, one needs the first  $j$  forward and  $(j-1)$  backward asymptotic GS vectors to obtain the upper left part of  $\mathbf{P}$  which is needed to compute the (forward) expansion coefficients of the first  $j$  CLVs.

### 6.3. Comments

To conclude, all static methods, be the direct intersection method of Eqs. (17-19) or the two refinements briefly discussed above, require the solutions of certain systems of linear homogeneous equations, which in turn depend on the forward and backward GS vectors. For direct subspace intersection, the first  $j$  forward and the last  $(N-j+1)$  backward GS vectors are required to compute the first  $j$  CLV, while only the first  $j$  forward and  $(j-1)$  backward vectors are needed by the refined algorithms - an improvement if one is interested in either the first or last CLVs.

However, in our opinion static approaches suffer from a number of disadvantages. First of all, one has to compute both forward and backward GS vectors, and thus is forced to perform twice vector orthonormalization at each trajectory point where CLVs are needed. Note that, as discussed in Section 4.2, vector orthonormalization is the computationally most demanding part of the dynamics for both short ranged and globally coupled large dynamical systems. Therefore, it is not a good idea to double the number of such orthonormalizations with respect to the plain dynamical algorithm of Section 4.

A second concern regards the solution of large systems of linear equations, which has to be performed by singular value decomposition (SVD) to attain a satisfactory numerical accuracy. However, SVD is more time consuming than back substitution by a factor 18, as it requires  $\sim 6m^3$  operations for an  $m \times m$  matrix [28].

Finally, for what regards memory requirements, only forward GS vectors need to be stored by the static algorithm. While this reduces the memory requirement to about

2/3 of what needed by the dynamical algorithm to store both  $\mathbf{G}$  and  $\mathbf{R}$  matrices, this memory advantage is lost whenever one is only interested in the angles between vectors, for which the dynamical algorithm only needs to store the upper triangular matrices  $\mathbf{R}$  and thus finds itself in a better position.

## 7. Angles between covariant vectors and subspaces

Covariant Lyapunov vectors provide direct information on the geometrical structure of tangent space. In particular, angles and (near)-tangencies between different CLVs or their associated subspaces can be used to characterize the dynamical properties of a chaotic dynamical system. In most applications of interest, this amounts to studying the distribution of such angles over the (ergodic) attractors. While angles are not invariant under a generic coordinate transformation, non-singular transformations do preserve zero and non-zero angles. Note also that CLVs corresponding to non-degenerate LEs may not become completely parallel along a given trajectory. Since they evolve via Eq. (28), it is clear that should at any point two vectors be completely parallel, they would stay the same along the entire trajectory, contradicting the non-degeneracy assumption. However, trajectories can pass arbitrarily close to such tangent points, resulting arbitrarily small angles. Therefore, relevant physical information has to be encoded in the way the angles probability distributions approaches the null angle.

In Refs. [16], angles between CLVs have been studied to show that the tangent space of generic spatially extended dissipative systems, such as the Kuramoto-Sivashinsky equation or the complex Ginzburg-Landau equation, is split into two decoupled subspaces. One comprises a finite number of frequently “entangled” CLVs, or *physical* modes, which carry all the relevant information of the trajectory. A second residual set is composed of strongly decaying *spurious* modes which are transversal to the “physical” manifold and themselves organized in mutually transversal subspaces. The number of physical modes, which is extensive in the system total degrees of freedom  $N$ , can be interpreted as the number of *effective* degrees of freedom needed to faithfully describe the chaotic dynamics, leading to the conjecture that the physical modes constitute a local approximation of the inertial manifold [35].

Another interesting issue regards the degree of non-hyperbolicity of a dynamical system. In hyperbolic systems [36, 37], there exists a direct sum decomposition of the tangent space  $\mathcal{T}_x\mathcal{M}$  at each point  $x$  into three *invariant* subspaces

$$\mathcal{T}_x\mathcal{M} = \mathbf{E}_x^u \oplus \mathbf{E}_x^s \oplus \mathbf{E}_x^0. \quad (70)$$

The unstable subspace  $\mathbf{E}_x^u$  is spanned by the CLVs associated to positive LEs, so that each vector  $\mathbf{u} \in \mathbf{E}_x^u$  is exponentially contracted backward in time. Similarly, the stable subspace  $\mathbf{E}_x^s$  is spanned by the CLVs associated to negative LEs and any  $\mathbf{u} \in \mathbf{E}_x^s$  is exponentially contracted forward in time. Finally,  $\mathbf{E}_x^0$  is associated to the null LEs and their corresponding covariant subspace<sup>§§</sup>. In particular, hyperbolicity implies that the stable and unstable subspaces are nowhere tangent. Dynamics on a hyperbolic attractor is structurally stable, i.e., is insensitive to variations of parameters. It manifests strong stochastic properties which allow for detailed theoretical analysis and useful results such as the shadowing lemma [38]. Violations of hyperbolicity may manifest themselves either as a change along the attractor of the number of

<sup>§§</sup>It is non-empty only for continuous time flows and/or in the presence of conservation laws which reduce the attractor dimensionality.

stable and unstable directions (*unstable dimension variability*) or through the presence of *homoclinic tangencies*, i.e. points where the stable and unstable manifolds are mutually tangent.

The fact that the overwhelming majority of dynamical systems of any practical use are not hyperbolic may seem a serious drawback, but the “chaotic” hypothesis [25] assumes that generic chaotic systems, in spite of violations of hyperbolicity, can be treated as essentially hyperbolic ones, allowing one to extend a number of useful results to non-hyperbolic chaotic systems. In particular, it implies ergodicity of the attractor and the existence of well-defined time averages with a probability distribution satisfying a large deviation law.

While the chaotic hypothesis has an evident empirical success, as testified by the positive tests of the fluctuation theorem [39, 40], it is nevertheless of great interest to assess the degree of violation of hyperbolicity in a given chaotic system, especially in relation to its dynamical properties. In the following, we concentrate on homoclinic tangencies.

### 7.1. Subspace intersection

As it has been correctly pointed out in Ref. [13], in order to compute the angle between two linear subspaces it is not sufficient to compute all the angles between pairs of vectors taken from two bases spanning the two subspaces. Indeed one has to consider angles between arbitrary linear combinations of such vectors. As shown in [13], this can be taken into account by SVD. Consider two generic subspaces, respectively spanned by  $m_1$  and  $m_2$  different CLVs,  $m_1 + m_2 \leq m$ , with  $m$  being the tangent space dimension. Suppose  $m_2 \geq m_1$ . First of all, we organize the CLVs from the two subspaces in two different matrices:  $\mathbf{U}_1$  (of size  $m \times m_1$ ), whose columns contain the vectors from  $V_1$  and  $\mathbf{U}_2$  ( $m \times m_1$ ), with the vectors from  $V_2$ . Since angles can be equally computed in the phase space coordinate basis as well as in the GS basis, we choose the latter and consider the proper columns of the matrix  $\mathbf{C}$ .

We have to compute the QR factorization of both matrices,

$$\mathbf{U}_1 = \mathbf{Q}_1 \mathbf{R}_1 \quad , \quad \mathbf{U}_2 = \mathbf{Q}_2 \mathbf{R}_2 \quad (71)$$

and compose the  $m_1 \times m_2$  matrix  $\mathbf{Q}_2^T \mathbf{Q}_1$ . There are  $m_1$  principal angles  $\theta_n^{(i)} \in [0, \pi/2]$  between the two subspaces, and their cosines are given by the principal values  $s^{(i)}$  of  $\mathbf{Q}_1^T \mathbf{Q}_2$ , that is

$$s_n^{(i)} = \cos \theta_n^{(i)} \quad i = 1, 2, \dots, m_1 \quad (72)$$

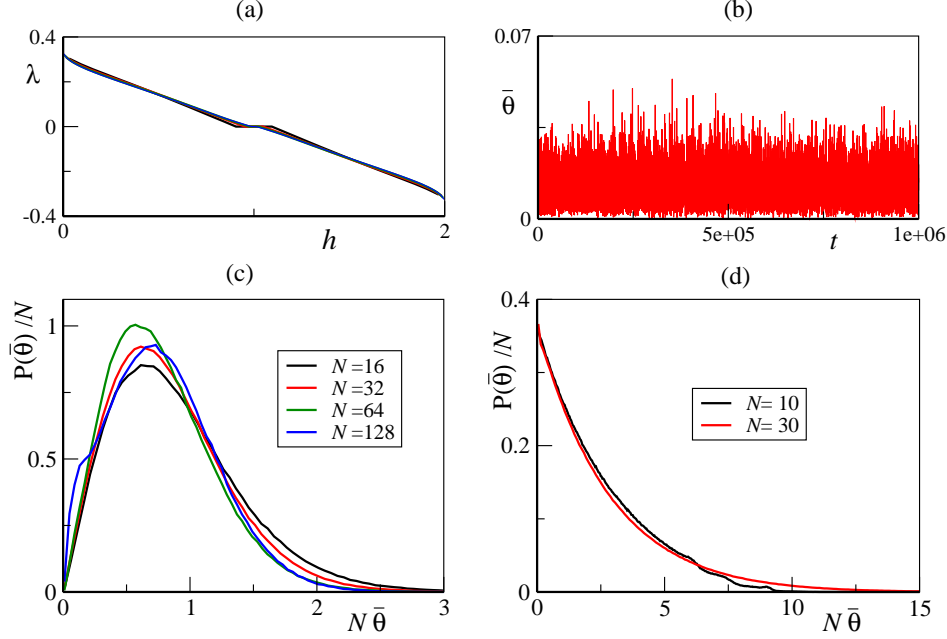
where we have explicated the time index  $n$ .

In our case we are interested in the intersection between the stable  $\mathbf{E}^s$  and the unstable  $\mathbf{E}^u$  manifolds. Note that since the unstable manifold  $\mathbf{E}^u$  is spanned by the first  $m_1$  CLVs, the matrix  $\mathbf{U}_1$  is upper triangular, and its corresponding orthogonal matrix  $\mathbf{Q}_1$  is just the identity matrix, thus simplifying the calculations of principal angles. In particular, we are interested in the minimum angle  $\bar{\theta}_n = \min_i \theta_n^{(i)}$ .

### 7.2. Numerical examples

In this section, we present two numerical examples: a Hamiltonian system, the FPU chain (59) with  $N$  oscillators, and a dissipative one, a chain of  $N$  Hénon maps (47). In the Hamiltonian case we have  $m_1 = m_2 = N - 2$ , since there are 4 null LEs in one dimensional FPU, corresponding to momentum and energy conservation and the





**Figure 4.** (color online) FPU (at energy density  $e = 10$ ) and Hénon chains results. (a) FPU Lyapunov spectrum as a function of the rescaled index  $h = (i - 0.5)/N$  for different system sizes,  $N = 16, 32, 64, 128$ . (b) FPU Minimum angle  $\bar{\theta}$  timeseries for  $N = 64$ . (c) FPU probability distribution  $P(\bar{\theta})$  as a function of the rescaled minimum angle for different system sizes,  $N = 16, 32, 64, 128$ . (d) Minimum angle rescaled probability distribution for the Hénon chain at sizes  $N = 10$  and  $N = 30$ . Other parameters as in Fig. 2

associated symmetries. In the Hénon chain case, one has typically  $m_1 = m_2 = N$ , and there are no zero LEs. We have investigated the tangent-space geometrical structure by measuring the minimum angle between the stable and unstable manifolds along a trajectory which samples the system's ergodic measure.

We start discussing the FPU system at an energy density  $e = H/N = 10$ , which is known to be characterized by a well developed chaotic dynamics. Our numerical simulations employed periodic boundary conditions and a McLachlan-Atela integration algorithm [41], which is well-suited for Hamiltonian systems, and an integration step of  $\Delta t = 0.05$  (we have verified the stability of our results versus  $\Delta t$ ). We employed a transient (both forward and backward) of  $t_0 = N \cdot 10^5$  time units, and sampled the minimum angle  $\bar{\theta}_n$  between the stable and unstable manifold along  $10^6$  time units (our sampling and orthonormalization rate is one time unit).

We considered  $e = 10$  and different system sizes between  $N = 16$  and  $N = 128$ . The rescaled Lyapunov spectra for this system are shown in Fig. 4a, (they have been computed both along the forward dynamics with the Benettin *et al.* algorithm as well as from the CLVs expansion rates (28), as a check of CLV numerical convergence). The stationary time-series corresponding to the minimum angle  $\bar{\theta}_n$  between stable and unstable subspaces is plotted in Fig. 4b, where one can see that it is rather irregular. We find therefore convenient to reconstruct the probability distribution  $P(\bar{\theta})$  (PDF) of the minimal angles. (see Fig. 4c). By comparing the distribution for different system

sizes, one can conclude that the minimum angle scales as

$$\bar{\theta} \sim \frac{1}{N} \quad (73)$$

as testified by the relatively good data collapse of the various PDFs. This is somehow intuitive, considered that the number of principal angles is equal to the stable and unstable manifold dimension. Notice that the probability distributions vanish algebraically (roughly linearly) as  $\bar{\theta} \rightarrow 0$ , are characterized by a finite maximum and decay exponentially fast for large angles.

This behavior can be compared with results for a chain of Hénon maps. Also for this dissipative case, the minimum angle scales as  $1/N$ , but the shape of the PDF is rather different: it starts with a finite value for  $\bar{\theta} \rightarrow 0$  and decays exponentially as the angle is increased (see Fig. 4d). Note also that these latter PDFs are characterized by a much larger width.

One can certainly conclude that the dissipative system is characterized by stronger violations of hyperbolicity. It is, nevertheless, important to remark that a vanishing PDF when  $\bar{\theta} \rightarrow 0$  does not imply hyperbolicity. In hyperbolic systems, the distribution of minimum angles is *bounded away from zero*[9]. A vanishing  $P(\bar{\theta})$  implies that trajectories exist which pass arbitrarily close to a (zero measure) set of homoclinic tangencies¶¶. The different behavior of  $P(\bar{\theta})$  in zero reveals, however, important differences in the spatial structure of homoclinic tangencies. In particular, if  $P(\bar{\theta}) \sim \bar{\theta}^\alpha$  for  $\bar{\theta} \ll 1$ , the measure of points characterized by angles (between the stable and unstable subspace) smaller than a certain threshold  $\theta_0$  will scale as  $\theta_0^{\alpha+1}$ . The implications of this scaling law for the global dynamics are the object of current investigation.

## 8. Concluding remarks

Covariant Lyapunov vectors provide an intrinsic decomposition of tangent space that is invariant under time reversal and independent of the norm (CLVs coincide with the Floquet eigenvectors when computed along periodic orbits [42]). For this reason, whenever individual directions need be considered, CLVs have to be chosen over other vectors, such as the asymptotic GS vectors or the “singular” vectors [7], which represent a finite-time version of the former ones, used in ensemble forecast applications and which do depend on the norm. Indeed, it has been shown [9] that the orthonormalization procedure (23) introduces, spurious structures in the individual GS vectors which have nothing to do with the underlying dynamics.

It would be interesting to apply CLVs for a better control of the uncertainty in nonlinear models such as those used for weather forecast. While variational data-assimilation techniques can benefit from the generic knowledge of the system’s unstable manifold [43], it is conjectured that the most relevant instabilities which affect large scale structures in atmospheric models are not those ones associated to the fastest time scales (i.e., the largest LEs), but rather to slower instability modes, characterized by longer spatial wavelengths. It logically follows that the most important directions for data-assimilation coincide with those associated to small positive LEs, which can be individually accessed only by computing CLVs. One may therefore hope that the

¶¶Here we assume CLVs in  $\mathcal{T}_x\mathcal{M}$  are continuous with respect to  $x$ . A sufficient condition for the continuity of the asymptotic GS base, which in turns implies the continuity of CLVs is discussed in [27].

ability to access the single unstable directions associated to well-defined timescales (the inverse of the LE) is of practical use in optimal forecast methods.

Covariant Lyapunov vectors have a number of further potential applications in dynamical systems theory. They can be used to assess the hyperbolicity of the underlying dynamics [9, 13] and, as shown in the previous section, can also contribute to shed light on the spatial structure of tangent space. In this respect, it is worth mentioning that preliminary simulations performed in the FPU chain at lower energies (below the strong stochasticity threshold) reveal stronger violations of hyperbolicity. Moreover, CLV have been also employed to characterize the collective dynamics of large chaotic systems through their localization properties [15], thus allowing to establish a connection between microscopic evolution and the emergence of global properties.

Finally, one puzzling finding in Hamiltonian and symplectic systems regards the  $k^{-1}$  divergence displayed by the power spectrum of the spatial part of the covariant vectors associated to the smallest LEs [9]. Further work is required to clarify whether this property can be related to the so-called hydrodynamic Lyapunov modes [44] or whether this  $1/f$  behavior has any relation with actual dynamical properties of these systems.

## Acknowledgments

We warmly thank G. Morriss, P. Kuptsov, S.P. Kuznetsov, A. Pikovsky and M. Sala for enlightening discussions.

- [1] Shimada I and Nagashima T, 1979 *Prog. Theor. Phys.* **61** 1605.
- [2] Benettin G, Galgani L, Giorgilli A and Strelcyn J M, 1980 *Meccanica* **15** 9; Benettin G, Galgani L, Giorgilli A and Strelcyn J M, 1980 *Meccanica* **15** 21.
- [3] Oseledets V I, 1968 *Trudy MMO* **19** 179.
- [4] Eckmann J-P and Ruelle D, 1985 *Rev. Mod. Phys.* **57**, 617.
- [5] Ruelle D, 1979 *Publ. Math. IHES* **50**, 275.
- [6] Brown R, Bryant P, and Abarbanel H D I, 1991 *Phys. Rev. A* **43**, 2787.
- [7] Legras B, Vautard R, 1996 *Predictability vol I*, Seminar Proceedings, ECMWF, Reading, UK, pp. 143.
- [8] Politi A, Torcini A, Lepri S, 1998 *J. Phys. IV* **8**, Pr6 263.
- [9] Ginelli F *et al.*, 2007 *Phys. Rev. Lett.* **99**, 130601.
- [10] Wolfe C L and Samelson R M, 2007 *Tellus* **59A**, 355.
- [11] Kuptsov P V and Parlitz U, 2012 *J. Nonlinear Sci.* **22**, 727.
- [12] Szendro I G, Pazo D, Rodríguez M A, and López J M, 2007 *Phys. Rev. E* **76** 025202(R); Pazo D, Szendro I G, López J M, and Rodríguez, 2008 *Phys. Rev. E* **78** 016209.
- [13] Kuptsov P V, Kuznetsov S P, 2009 *Phys. Rev. E* **80**, 016205.
- [14] Yang H-l, Radons G, 2008 *Phys. Rev. Lett.* **100**, 024101.
- [15] Takeuchi K A, Ginelli F, Chaté H, 2009 *Phys. Rev. Lett.* **103** 154103.
- [16] Yang H-l, Takeuchi K A, Ginelli F, Chaté H, and Radons G, 2009 *Phys. Rev. Lett.* **102** 074102; Takeuchi K A, Yang H-l, Ginelli F, Radons G and Chaté H, 2011 *Phys. Rev. E* **84** 046214.
- [17] Bosetti H, Posch H A, 2010 *Chem. Phys.* **375**, 296; Bosetti H, Posch H A, Dellago C, 2010 *Phys. Rev. E* **82** 046218.
- [18] Truant D P, Morriss G P, 2011 *J. Stat. Mech.* P010114.
- [19] Ouchi K, Horita T, Yamada T, 2011 *Phys. Rev. E* **83** 046202.
- [20] Inubushi M, *et al.* 2012 *Phys. Rev. E* **85**, 016331.
- [21] Lopes S R, *et al.*, 2012 *Phys. Rev. E* **86**, 016216.
- [22] Yang H-l, Radons G, Kantz H, 2012 *Phys. Rev. Lett.* **109**, 244101.
- [23] Pesin Ya B, 1976 *Izv. Akad. Nauk SSSR, Ser Mat. Tom* **40**, 1261.
- [24] Pesin Ya B, (1977) *Russ. Math. Surv.* **32**, 55.
- [25] Ruelle D, 1980 *Ann. New York Acad. Sci.* **357**, 1; Gallavotti G, Cohen E G D, 1995 *Phys. Rev. Lett.* **74**, 2694; Gallavotti G, Cohen E G D, 1995 *J. Stat. Phys.* **80**, 931; Gallavotti G, 2006 *Chaos*, **16**, 043114.

- [26] Guckenheimer J, Holmes P, 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (New York: Springer).
- [27] Ershov S V and Potapov A B, 1998 *Physica D* **118**, 167.
- [28] G.H. Golub, C. F. Van Loan, 1996 *Matrix Computations*, 3rd edition (John Hopkins University Press).
- [29] Ruelle D, 1982 *Commun. Math. Phys.* **87**, 287; Livi R, Politi A, and Ruffo S, 1986 *J. Phys. A* **19**, 2033.
- [30] Takeuchi K A, *et al.*, 2011 *Phys. Rev. Lett.* **107** 124101.
- [31] Politi A and Torcini A, 1992 *Chaos* **2**, 293.
- [32] Touchette H, (2009). *Physics Reports* **478**, 1.
- [33] Crisanti A, Paladin G, and Vulpiani A, 1988 *J. Stat. Phys.* **53**, 583.
- [34] Pettini M *et al.*, 2005 *Chaos* **15**, 015106.
- [35] Yang H-l and Radons G, 2012 *Phys. Rev. Lett.* **108**, 154101.
- [36] Ruelle D, 1989 *Chaotic evolution and Strange attractors* (Cambridge: Cambridge University Press).
- [37] Barreira L and Pesin Ya B, 2007 *Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents* (Cambridge: Cambridge University Press).
- [38] Bowen R, 1970 *American Journal of Mathematics* **92**, 725.
- [39] Lepri S, Rondoni L, Benettin G, 2000 *J. Stat. Phys.* **99** 857.
- [40] Ciliberto S, Laroche C, 1998 *J. Phys. IV* **8** 215; Garnier N, Ciliberto S, 2005 *Phys. Rev. E* **71** 060101.
- [41] McLachlan R I, Atela P, 1992 *Nonlinearity* **5** 541.
- [42] Trevisan A, Pancotti F, 1998 *J. Atmos. Sci.* **55** 390.
- [43] Trevisan A, D’Isidoro M, and Talagrand O, 2010 *Quart. J. Roy. Meteor. Soc.* **136** 487; Trevisan A, Palatella L, 2011 *Int. J. of Bifurcation and Chaos* **21** 3389.
- [44] Eckmann J-P, Forster Ch, Posch H A, and Zabey E, 2005 *J. Stat. Phys.* **118** 813.